AN INVESTIGATION INTO THE EFFECT OF DIAGRAMS ON PERFORMANCE IN ORDINARY GRADE PHYSICS QUESTIONS.

NEIL R. SHORT.

Submitted in part fulfillment
of the requirements for the
Degree of Master of Education.

August 1980

S U M M A R Y

A Picture is worth a

thousand words

Chinese Proverb

This research attempts to examine the effect of the inclusion of diagrams in Ordinary Grade standard Physics test papers of the extended answer variety.

CHAPTER 1 contains a general outline of the main factors in the test design, the sample, validity, the hypothesis itself and the type of statistical test to be used.

CHAPTER 2 outlines a proposed hierarchy of diagram types, and goes on to consider in detail how the factors mentioned in Chapter 1 relate to the experiment.

CHAPTER 3 details the results of the tests applied to the sample, and the result of the opinion test conducted in one of the schools.

CHAPTER 4. contains conclusions and recommendations for further research into this subject area.

CONTENTS

Acknowledgements

Thanks are due to Mr John Timmins, the Headmasters and Principal Teachers of Physics in the schools visited, for their help in the selection and testing of the pupils. Mention must also be made of Messrs Pillans and Wilson Ltd. for permission to reproduce several of their illustrated test questions. Finally, thanks go to Bob Neill of my School Physics Department for igniting the original spark of this enquiry, and particularly to my supervisor Alastair Pollitt for fuelling the fire and indeed raking the ashes when required.

N.R.S.

CHAPTER 1

By a small sample we may
judge of the whole piece.

CERVANTES.

The author has been involved in the production of a substantial number of Physics test papers in recent years, both as a teacher in the making up of material for internal assessment, and externally, in the writing of several papers for Messrs. Pillans and Wilson Ltd., who produce a range of papers frequently used by schools in preliminary examinations for the Scottish Certificate of Education.

In the case of the external papers in particular, the necessity of working with an "artist" to produce diagrams, has brought into focus the question of whether such diagrams have any measurable effect on the performance of the candidates.

Work of recent years into the influence of imagery in the learning process as published for example by Varley, Levin, Severson, and Wolff (1974, 1) suggests that drawing can facilitate learning.

Little research seems to have been conducted, however, into the influence of pictorial content in test items, there being no information available in the more likely sources of the literature (A1 - A5)

In order to shed some light on this question, the following work has been undertaken to determine the effect, if any, of diagrammatic content on the performance of Ordinary Grade Physics candidates in an extended answer test.

The tests were administered to a sample of 184 fourth year 'O' Grade pupils at a time after their preliminary examinations had taken place, and before their 'O' Grade examination itself in May of 1980.

SAMPLING

It is rarely easy to conduct an experiment involving an entire population; the researcher is usually forced by limited resources to select a sample which is expected to be representative of that population. Using the statistic obtained from the sample, he hopes to generalise. As Moser (2) points out, one of the decisive factors in sample design is the nature of the Sampling Frame, that is, the list or index of the population available to the researcher.

Desirable characteristics in a sampling frame are:-

(a) Adequacy - that the frame covers the entire population concerned in the enquiry.

(b) Completeness - that all members of the chosen population are listed.

(c) Non-duplication of entries.

(d) Accuracy - the removal of "non-existent" entries e.g. candidates who have withdrawn; the addition of late entrants.

Given that an effective sampling frame has been constructed, the method in which the sample is drawn has an important bearing on the outcome of the experiment. Various sampling methods are available.

(a) Simple Random Sampling

The essence of this method is that each member of the population has an equal chance of selection. Thus, given a complete sampling frame, simple random sampling protects against selection bias.

(b) Stratified Sampling

Here, the population is divided into strata e.g on the basis of school size or geographical location. Random sampling then occurs from each stratum. This technique can increase the precision of an experiment as stratification effectively reduces standard error.

(c) Matched Sampling

This type of sample involves the matching of pairs in the sample to produce two "sub-samples" which are comparable on some factor. In this research, a sample of diagram test candidates is matched or paired with so-called non-diagram test candidates on the basis of a pretest, the Ordinary Grade preliminary examinations. This method again reduces the standard error making for a more sensitive application of the statistical test, in this case student's t test for matched samples.

VALIDITY

Generalizability, as Pilliner (3) points out, is also a function of the internal and external validity of a test. For a test to be internally valid, its results must be credible "within its own confines", that is, the result must be a function of some chosen treatment (often referred to as the independent variable) and not of some extraneous variable which has not been allowed for or included in the test.

External validity relates to the applicability of the sample result to other similar situations or groups. Internal and external validity are related in that internal validity is a necessary (but not sufficient) condition for external validity. The converse, however, is not true.

A certain incompatability exists also between the two forms of validity. The tighter the constraints applied to a sample to ensure internal validity, the less likely is the result to apply generally and thus achieve external validity.

Campbell and Stanley (4) mention several factors which affect the internal and external validity of a test. Among those having an influence on internal validity are:-

(a) History

Clandestine learning events which occur between a pre-test and a post-test in addition to the treatment under study.

(b) Maturation

The effects of the passage of time between pre-test and post-test e.g. on physiological condition, attitudes etc.

(c) Testing

The effect of a pre-test itself on results in the post-test.

(d) Regression

The effect obtained when a sample is selected on the basis of some extreme variable e.g. the highest 5% of I.Q. Any group selected on the basis of an extreme score in one test will have a mean which on later re-test will tend to "regress" towards the population mean. Thus, high performers in a pre-test may later appear to have done less well in a post-test. Similarly, poor performers may seem to be showing

an improvement. Such differences may be falsely attributed to the effect of the independent variable.

(e) Bias

An undetected, pre-existing difference between an experimental and a control group may be assumed wrongly to have been caused by the treatment applied to the experimental group.

(f) Experimental Mortality

The loss of a disproportionate number of able pupils for instance from one of two groups will alter the outcome of a comparison test.

The external validity of a test can be affected by:-

(g) The Interaction Effect of Testing

A sample of pupils may learn from a pre-test and therefore no longer will be representative of the un-pretested population.

(h) The Effect of Selection Bias

(i) The Effect of the Experimental Arrangements

This effect sometimes referred to as the Hawthorne Effect, can occur in groups who receive special attention, or a novel form of tuition. They may perform better than a control group remaining in a normal environment.

The "improvement" may be falsely attributed to the effect of the independent variable.

Having selected a sample and adopted an experimental design to combat the problems associated with validity, it is necessary for the researcher to formulate a hypothesis to which he may apply an appropriate statistical test.

THE HYPOTHESIS

In this research, a Null Hypothesis will be adopted, that is, one which states that no statistically significant difference in performance will occur between the two groups under test, (in this instance a group sitting a diagram test and a group sitting a non-diagram-test). This is equivalent to the statement that the two groups come from the same population as regards performance i.e. their mean scores belong to the same population distribution of scores.

In order to arrive at a decision regarding the validity of the Null Hypothesis, a statistical test is applied to the difference in means of the two groups (Student's t test). The value of t obtained indicates the probability that the two means come from the same population. A pre-determined probability level, in this research, the 5% level then is used as a criterion for acceptance or rejection of the Null Hypothesis.

t is calculated on the assumptions that the two samples come from populations having the same variance, and that those populations are normally distributed.

t is obtained thus -

$$t = \frac{\bar{X}d}{\sqrt{\frac{\Sigma x_d^2}{N\,(N-1)}}}$$

where $\bar{X}d$ is the mean of the N differences between matched pair scores and xd is the deviation of a single difference from the mean $\bar{X}d$ (Guilford and Fruchter, 5)

The specific calculations appear in Appendix (p 44)

As the value of t only indicates a probability, it is possible to commit two types of error in accepting or rejecting the Null Hypothesis.

Type 1 Error.

Rejection of a Null Hypothesis which should be accepted

The probability of a Type 1 error is indicated by the level of significance chosen with regard to the Null Hypothesis. In this research, a 5% level has been chosen. There is thus a five in one hundred chance that a decision to reject the Null Hypothesis will be wrong, given a t value which is termed significant.

Type II Error

Acceptance of a Null Hypothesis which should be rejected.

The probability of this type of error is often very difficult to assess - Johnson (6) suggests that it may be possible to propose a second, false hypothesis and then estimate the probability of its acceptance.

It may be seen from the above that any reduction in the probability of a Type 1 error by reduction of significance level e.g. from 5% to 2% will result in an increase in the probability of a Type II error occurring.

CHAPTER 2

THE TESTS

Diagrams are a common and accepted feature of Physics examinations today. Abrief survey of questions appearing in Ordinary Grade extended answers papers has suggested the use of a simple hierarchy of question types:-

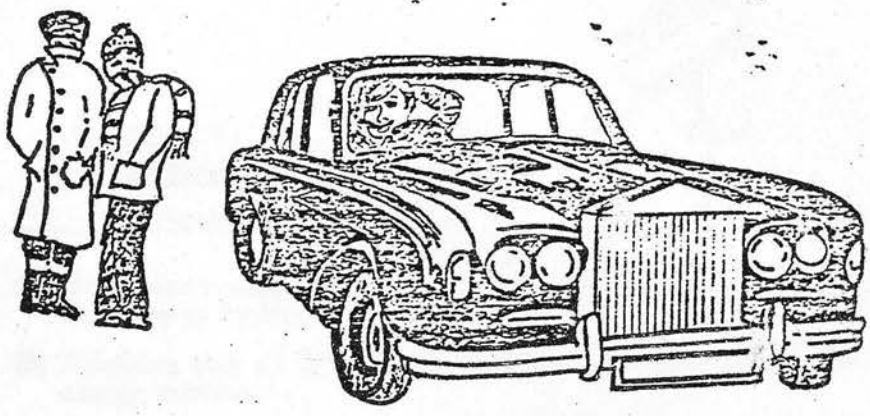
Type 1

Questions containing no diagrams.

Type 2

Those in which a diagram contains no information necessary for the answering of the question (fig. 1, overleaf). This type can have no direct cognitive effect on the candidate, but it might be argued that some reaction is produced e.g. amusement, relaxation which in itself may affect performance in the test.

4. Larry's luxury limousine has an air-conditioning system which maintains a constant temperature of 20°C inside the car regardless of the temperature of the air outside. Every 50 seconds, 5 m^3 of air passes into the interior and out through vents at the rear window. Since the volume of the car interior is 5 m^3, the system provides one complete air change in 50 seconds.



(*a*) Calculate the heat energy required for one complete air change if the system is operating while the outside temperature is 15°C. [3]

(*b*) What is the effective power rating of the system when it is operating as described in (*a*)? [2]

(*c*) In practice, the system would require a higher power rating. Explain this. [2]

(*d*) The limousine also has a small 25 W heater for de-icing the windscreen on very cold days. If 10 g of ice cover the windscreen on a morning when the outside temperature is −10°C, calculate the minimum time needed to melt all of the ice on the windscreen. [3]

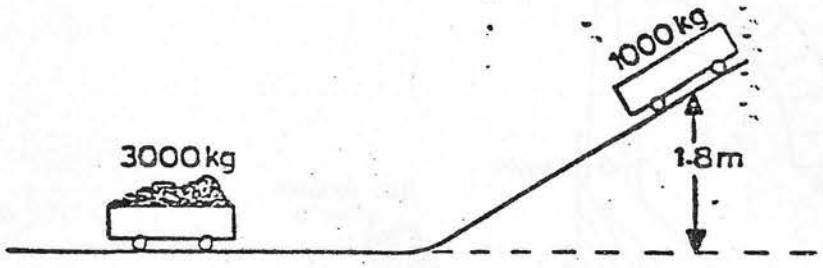
Specific heat capacity of air $= 1{\cdot}0 \times 10^3$ J kg^{-1} K^{-1}.
Density of air $= 1{\cdot}29$ kg m^{-3}.
Specific heat capacity of ice $= 2{\cdot}1 \times 10^3$ J kg^{-1} K^{-1}.
Specific latent heat of fusion of ice $= 3{\cdot}34 \times 10^5$ J kg^{-1}.

Type 3

Questions in which the diagram provides some duplication or reinforcement of information already in written form in the question. This type of question appears commonly in Scottish Certificate of Education Ordinary Grade Papers in Physics. The Examination Board has no official policy on the inclusion of diagrams in extended answer questions, but informally considers them to be useful as reinforcers of information already provided in written form.

(see overleaf)

3. In a railway goods yard, an empty wagon of mass 1,000 kg has brake failure. Starting from rest, it moves down a slope and then collides with a loaded wagon, of total mass 3,000 kg, which is standing on the level at the bottom of the slope.



(*a*) How much potential energy did the empty wagon have when it was at the top of the slope? [2]

(*b*) Assuming that all this potential energy is converted into kinetic energy, calculate:—

(i) the velocity of the empty wagon when it reaches the bottom of the slope;

(ii) the momentum of the empty wagon when it reaches the bottom of the slope. [3]

(*c*) After colliding, the loaded wagon moves off at 2·5 m s^{-1} to the left. Explain in detail what happens to the empty truck, by considering conservation of momentum. [3]

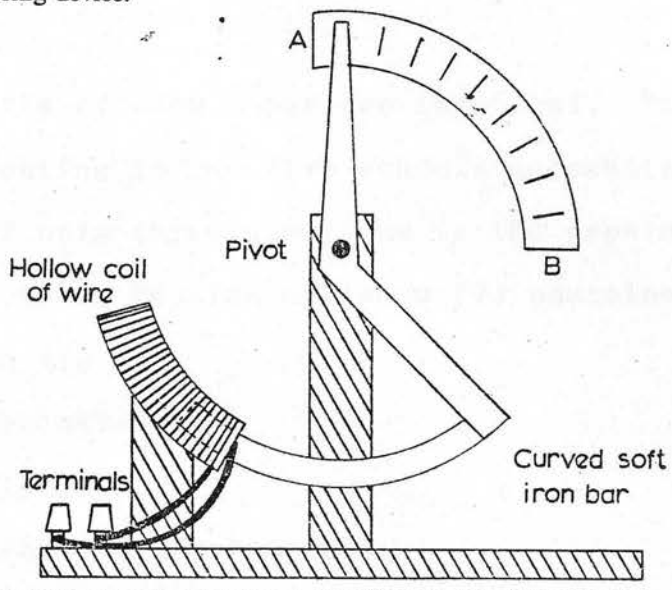
(*d*) How much kinetic energy is lost during this impact? What happens to this 'lost' energy? [2]

Type 4

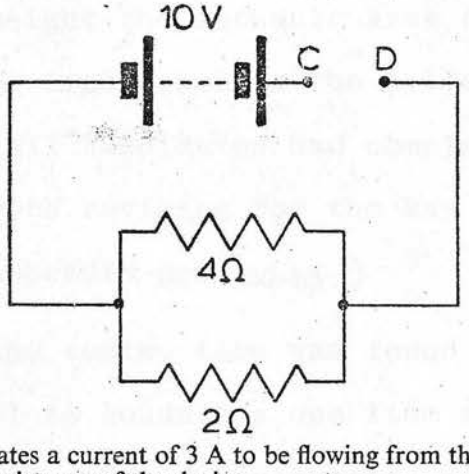
Questions in which information is provided exclusively in diagrammatic form. Removal of the diagram from the question would render the question unanswerable.

(see overleaf)

7. The diagram illustrates a student's attempt to construct a current measuring device.



(*a*) Explain as fully as you can exactly how this device works. [3]
(*b*) Would the device still work if the soft iron bar were replaced by (i) a brass bar; (ii) a glass bar?
Give a reason for your answer. [2]
(*c*) Explain what would happen if the current through the device were reversed. [2]
(*d*) The device is now connected into the gap CD in the following circuit.



If it indicates a current of 3 A to be flowing from the battery, calculate the resistance of the device. [3]

This research has been designed to look for any effect produced by a diagram in Type 3 questions. To this end, a "control" paper has been produced, based very closely on previous 'O' Grade questions of Type 3// . The second or "experimental" paper contains the same information but in Type 1. format.

The question parts of each paper are identical. Time available for testing in the five schools necessitated the inclusion of only three questions in the papers. The present 'O' Grade Physics Syllabus (7) contains four main topic areas viz.

(1) Mechanics

(2) Electricity

(3) Heat and the Gas Laws

(4) Waves

Type 3 questions appear largely in areas (1) and (3). It was therefore decided to include those areas in the paper and to weight the mechanic area (2 questions) reflecting its greater importance in the syllabus. At the time of testing, all candidates had completed the syllabus topics and were revising for the May examination. (The tests appear in Appendix pps 50-55:)

In addition to the tests, time was found at Craigmount High School to conduct a one item attitude or opinion test. After a brief explanation of the purpose of the research, each candidate answered the following question at the end of his test paper:-

"As compared with non-diagram questions, having diagrams in my question paper makes the test:- "

(A) Very much easier

(B) A little easier

(C) No different

(D) A little more difficult

(E) Very much more difficult

The results and discussions appear on pp 24-26

THE SAMPLE

For reasons of Geography and time available the sample obtained in this research does not completely meet the requirements stated in the previous chapter. It does, however, include both rural and urban comprehensive schools and a grant-aided school. Within three of the schools, Craigmount, Kelso and Selkirk it was possible to obtain the entire set of fourth year Ordinary Grade Physics candidates, a total of 128 pupils. Thus, a suitably wide range of candidate ability appears to have been sampled.

TABLE 1

The Sample

School Type	School	Sample Size
City Comprehensive	Craigmount	70
	St. Thomas of Aquins	14
Rural Comprehensive	Kelso	38
	Selkirk	20
City Grant-Aided	George Heriot's	42

It will be argued, therefore although some bias may well exist in the sample, it is nevertheless moderately representative of the population and a cautious generalization of results might still be of value.

SELECTION OF THE COMPARISON GROUPS

Two possible methods of selection of experimental and control groups were available, viz. random selection and pairing on the basis of some pretest information. The latter design can yield greater efficiency given that the pre-test results correlate sufficiently highly with those obtained in the latest test. In this case, information was available from the Schools' Ordinary Grade "prelim" results on the performance of all of the candidates. It was therefore decided to adopt a procedure of selection of matched pairs to form experimental and control groups. In three of the schools, pairing was made on the basis of actual test scores, while in the other two, rank order of merit was used, the outcome being the same.

It seems to be a reasonable assumption that a high correlation will exist between "prelim" scores and those scores from the research test, particularly as the research test was itself based very closely on previous Ordinary Grade questions. However, as a check, a correlation coefficient was obtained for sixteen raw scores chosen at

random from the Craigmount High School marks. The value of r obtained, 0.73 seems to support this assumption. (p 45)

VALIDITY AND THE TEST DESIGN

A large number of the problems associated with external and internal validity are removed by the process of matching of the test groups. These are the factors associated with History, Bias, Maturation and Regression towards the Mean. The problem of bias may extend beyond the setting up of the matched groups, in the form of marker bias, that is, different markers may have different standards of assessing, or a single marker may influence results through a preference (even unconsciously) for one group. In this research, a single marker worked to a detailed marking scheme and was largely unaware of the group (diagram or non-diagram) from which a given paper had come. The effect of prior tests in Physics on validity however, deserves some extra mention.

The Pre-Test and Prior Testing in General

A danger exists, when questions in the research diagram tests are based on previous Ordinary Grade questions, that pupils may benefit from recall. Recognition may be more likely in the group sitting the diagram test as the three questions therein bear a much greater resemblance to past questions than do those in the

non-diagram test. On the assumption then that many pupils will have had access to past papers during revision, any result showing a significantly greater mean score in the diagram test group would have to be treated with caution.

As regards the interaction effect of testing and its effect on external validity, it seems safe to argue that the vast majority of the population experience a pre-test in the form of a preliminary examination. Access to past papers is also widespread. Finally, any effect due to the experimental arrangements should be discountable, as the candidates were unaware of the nature of the test, it being presented as a revision test. Only at the close of the test were they informed that two different papers had been administered.

THE STATISTICAL TESTS

Two tests were performed on the available data:-

(1) Student's t test was applied to the mean difference of scores of the entire sample on the hypothesis that this mean difference would not be significant.

(2) In order to investigate further a suspected interaction between ability and test type, a correlation coefficient was obtained and a Fisher's Z calculated to determine whether the correlation was significant.

Finally, the opinions of the Craigmount High School pupils on the effect of diagrams in the test were tabulated.

CHAPTER 3

A reasonable probability is
the only certainty.

Edgar Watson Howe.

The results of educational research rarely provide certainties. In that respect, this work is no exception. It will be shown, however, that some light has been cast on the problem discussed at the beginning of this work, and perhaps more importantly, some pointers to renewed investigations have been suggested.

THE t-TEST RESULT

The value of t obtained in the first test was 1.152. On the predetermined 5% level adopted for the Null Hypothesis, this value falls considerably short of the necessary figure of approximately 1.96. Thus we are able to conclude that no significant difference exists between the means of the diagram and non-diagram tests. Any intuitive idea that diagrams will assist candidates is certainly not supported in this test;

indeed, the mean difference is a positive 0.554 showing a slightly better non-diagram candidate performance.

INTERACTION OF TEST TYPE AND ABILITY

An initial investigation was conducted into a possible interaction between the effect of diagrams and ability. The sample pairs were divided into three groups plotted on a graph. The suspicion of a trend indicated by the graph was further investigated by obtaining a correlation coefficient between effect of diagrams and ability. (See (pps 49,50 for Graph).

The Effect of Diagrams

The quantity chosen to reflect this variable was (Diagram Score - Non-Diagram Score) for each of the 92 matched pairs. This seems an appropriate and logical choice, as the difference in score will reflect the effect of the test type, given that the matched pairs have the same ability, (the assumption on which the test design is based). The assumption was checked by obtaining a correlation coefficient between diagram and non-diagram test scores. The value of 0.965 obtained would seem to confirm that the original matching was suitable precise.

Ability

The combination of scores chosen to act as a measure of ability was the sum of diagram and non-diagram test scores. Strictly, the mean score of a matched pair would be the proper choice, but for computational convenience 10 X (Diagram Score + non-diagram score) was chosen, the result being ultimately the same. The choice was made on the basis that diagram and non-diagram scores were totally independent.

Using the scores in the above fashion, a correlation coefficient, r was obtained thus:-

$$r = \frac{\text{Covariance } (D+N)\ (D-N)}{\sqrt{\text{Variance } (D+N) \text{ Variance } (D-N)}}$$

The calculated value of r was 0.151

The value was then examined for significance by conversion to Fisher's Z and finding the deviation of Z from Zero.

$$\text{deviation} = \frac{\frac{1}{2} \log_e \frac{1 + r}{1 - r}}{\sqrt{\frac{1}{n - 3}}}$$

The value of the deviation obtained 1.44, was short of the value 1.96 required for significance at the 5% level. The deviation however, should not in the author's opinion be rejected out of hand and further investigation here will be recommended. (The calculat-:ions are to be found on pps 46, 47)

The final result obtained in the investigation relates to the attitude or opinion of the Craigmount High School pupils on the advantage if any of the inclusion of a diagram in a Physics test question. The results are shown in Table 2 below.

TABLE 2

Question:-	Having diagrams in a test paper makes the test	
Response	D. Candidates %	N.D. Candidates %
Very much easier	10.5	10.5
A little easier	39.5	35.0
No different	0	4.5
A little more difficult	0	0
Very much more difficult	0	0

It is clear from the table that most pupils think that diagrams help performance. Arguably then, they would be happier to see a diagram in a Physics Test question. This result may be the most important to come from the present research for teachers who are sensitive to the needs of their pupils.

In the absence of any clear proof of the effect of diagrams on performance, the inclusion of diagrams particularly early in a paper may produce a relaxing or reassuring effect on some of the pupils.

CHAPTER 4

The original purpose of the research was to determine whether the inclusion of diagrams of a type referred to as Type III earlier in this report, had any measurable effect on the performance of candidates in Physics examinations, specifically at 'O' Grade standard.

The result has been in statistical terms non-significant. In other words, the mean scores obtained by the diagram and non- diagram groups may be said to have come from the same population; the differences between the means are likely to have occurred by chance.

This "global" result should however, not be allowed to obscure certain detail results and observations which may ultimately have some relevance to those writing questions for Physics examinations:-

(1) Children themselves seem to be in little doubt that diagrams are of assistance when they are included in a test paper. The recognition of familiar experimental situations or pieces of apparatus in diagrams may have a confidence boosting effect, although whether this is translated into improved performance is far

from proven.

(2) Four candidates in the non-diagram group made a detail error in their responses to question 1b (p 53). They answered "Point X" instead of "Point Y" when it was clear from their responses later in that question that they understood the physics of the situation properly. A diagram would probably have helped here. This situation in some ways a relatively trivial one, highlights a problem which will be mentioned again later, namely that within the mass of the data may be hidden distinct effects on the individual candidate. Some candidates, for instance, may react strongly to the presence or absence of a diagram.

(3) The suspicion of a trend towards relatively lower performance in diagram tests by the poorer candidates was not confirmed in the test, but the matter clearly merits further investigation. One factor emerges from an examination of the test scripts of the lower third ability candidates, that is the greater proportion of diagrams drawn by non-diagram test pupils. These pupils drew a total of 24 diagrams as compared with 5 drawn by their diagram test partners. Has the action

of drawing itself helped poorer pupils to recall or visualise an experimental situation? The present research cannot answer this difficult question, but taken along with the other factors mentioned in (1) and (2) it may point in the direction of further work.

TWO POINTERS TO FURTHER RESEARCH

(1) A Further Examination Using Extended Answer Questions

The interaction of ability and the effect of diagrams could usefully be investigated further using a technique improved in the following ways:-

(a) Greater validity in relation to the Ordinary Grade Examination would be achieved by extending the test lengths from three to five questions and correspondingly increasing the syllabus area tested.

(b) The addition of information from another matched sample of pupils would help clarify the interaction effect. It would be necessary to obtain a Fisher's Z value and combine it with that obtained in this research to produce a mean value. Calculation suggests that a total sample size of 170 pairs would be the minimum required to show the presence of a significant effect.

(c) Further consideration of the hierarchy of diagram types proposed in Chapter 2 would suggest a subdivision of Type III into two sub-categories

(i) Diagrams of familiar arrangements previously demonstrated by a teacher, or relating to an experiment performed by the pupils themselves.

(ii) Diagrams describing novel and unfamiliar experimental apparatus.

2. A New Investigation Using Multiple Choice Items

Further work into the effect of diagrams could be pursued using pre-validated multiple choice items, such as are available in the Examination Board's item Bank. A forty item test could greatly increase the importance of any effect present, but not detected, in the "three item" extended answer test used in this research. The technique could be used to examine two further aspects of the problem.

(a) The interaction of the independent variable and the Bloom Category of the item. (as modified in the Scottish Education Department's Curriculum Paper 1969 - 8). While allocation of a question to a given Bloom category is not always precise, it is more easily effected in a multiple choice test.

(b) The Construction of a Type of Split-half multiple choice test. Such a test would contain twenty items of diagram type and a further parallel twenty non-diagram items. An examination of the standard deviation of differences per child could shed light on individual performance and individual reactions to the presence or absence of diagrams.

This work was undertaken in an attempt to answer a question arising from the author's own work as a teacher. The verdict so far must remain "not proven", but clearly, more work can be done in this area to produce items which are a true test of childrens' Physics abilities.

****

APPENDIX

Explanations of some Symbols and Abbreviations

$\bar{x}_D$ The mean score of a Diagram Test Group.

d The difference between the scores of a matched pair of candidates.

Σd The sum of differences.

N The number of matched pairs.

$\bar{X}_d$ The mean difference i.e. $\frac{\Sigma d}{N}$

Σx_d^2 The corrected sum of squares of the differences i.e. $\Sigma d^2 - \frac{(\Sigma d)^2}{N}$

df The number of degrees of freedom.

TABLE 3

RAW TEST SCORES

FOR

CRAIGMOUNT HIGH SCHOOL IN RANK ORDER

Diagram Test				Non Diagram Test			
Q1 /10	Q2 /10	Q3 /10	Total /30	Q1 /10	Q2 /10	Q3 /10	Total /30
9½	8	9	26½	9½	8	9	26½
9	9	8	26	8	8	9½	25½
9	8	8½	25½	9	8	8	25
8½	8	8	24½	8	8	8	24
8½	7½	7	23	8½	8	6½	23
8½	6½	8	23	8	6	8½	22½
5½	9	8½	23	9½	7	6	22½
8	8	6½	22½	6	8	8	22
9	6	7	22	6½	7½	7	21
8½	4½	8	21	6½	8	6½	21
8½	5½	7	21	6½	7½	6½	20½
6	7	8	21	8½	6	5½	20
6½	7	7	20½	8	4	8	20
8½	6	6	20½	7½	6	6½	20
7½	6	6½	20	8½	5½	6	20
8½	5	6	19½	4½	6	8	18½
8½	4	6½	19	6½	5	6½	18
5½	6	6	17½	6	4½	6	16½
5	5½	6	16½	5½	6	4½	16
7	5	4½	16½	7½	4	4½	16
5½	4½	6	16	3½	5	7	15½
5½	4	5	14½	5½	5	5	15½

TABLE 3 CONTINUED

Q1	Q2	Q3	Total	Q1	Q2	Q3	Total
5½	7	2	14½	4½	7	4	15½
5½	6	2	13½	5½	5	4	14½
2½	4	5	11½	3	4½	7	14½
4	3	3	10	5½	4½	4	14
½	2½	6	9	4½	3	6	13½
1	3	2	6	5½	5	3	13½
2½	2	1½	6	6	6	1	13
8½	7½	7	23	1½	4½	6	12
9½	5	7½	22	5½	4	2½	12
6	6	9	21	4	4	3	11
6½	6	5	17½	10	8	7	25
5½	6	6	17½	8	7	6	21
4½	4	5	13½	8½	5½	2	16

TABLE 4

RAW TEST SCORES FOR ST. THOMAS OF AQUINS

Diagram Test				Non-Diagram Test			
Q1	Q2	Q3	Total	Q1	Q2	Q3	Total
10	8	9	27	9½	8	10	27½
8½	8	8	24½	8½	9	8	25½
8½	8½	6	23	8½	6	8	22½
8½	7½	3	19	4½	8½	7	20
5½	5	6	16½	7½	6	6½	20
7	6	3	16	8	6	5½	19½
6	6	6	18	5½	5	8	18½

TABLE 5

RAW TEST SCORES FOR KELSO HIGH SCHOOL

Diagram Test				Non-Diagram Test			
Q1	Q2	Q3	Total	Q1	Q2	Q3	Total
10	7	8½	24½	9	9	9½	27½
9½	6	7½	23	7	7½	6	20½
6	9	7	22	4½	7½	4½	16½
5½	8	8½	22	5½	5½	5½	16½
5½	8	8	21½	3½	6½	6½	16½
6½	6½	8	21	4½	5	6½	16
7	6	6½	19½	8½	4½	3	16
5	8½	4½	18	6½	5	3½	15
4½	6½	6½	17½	5½	5½	2½	11
7	6½	4	17½	4	3½	5	12½
7	6	4	17	5½	4½	2½	12½
4½	5	3½	13	4	3	4½	11½
6	2½	2½	11	4½	4½	2	11
5½	2	3	10½	4	3	2½	9½
2	3½	3½	9	6½	5	8	19½
4	2½	2	8½	4½	6½	7½	18½
4	2	½	6½	5½	5½	7½	18½
2½	2	1	5½	5	7	6	18
4½	2	2½	9	4	5½	5	14½

TABLE 6

RAW TEST SCORES FOR SELKIRK HIGH SCHOOL

Diagram Test				Non- Diagram Test			
Q1	Q2	Q3	Total	Q1	Q2	Q3	Total
9	7	8½	24½	8½	8½	9	26
8	6½	7½	22	8½	7	9½	25
3	8½	8½	20	5½	9½	8	23
5½	6½	6½	18½	5½	5½	7	18
5½	5½	3½	14½	6½	6	5	17½
4	5½	4	13½	8½	4	1	13½
5	4	3	12	5½	4	1½	11
4½	2½	4	11	4	5½	0	9½
3½	5½	1	10	2½	2	1½	6
6	6½	6½	19	7½	6	3	16½

TABLE 7

RAW TEST SCORES
FOR
GEORGE HERIOT'S SCHOOL

Diagram Test				Non-Diagram Test			
Q1	Q2	Q3	Total	Q1	Q2	Q3	Total
6	4½	4	14½	3½	4½	4½	12½
4½	2½	1	8	8½	7	6	21½
5½	3½	4	13	3	7	5	15
9	6	2	17	3	6	6	15
5½	3	1½	10	6	6	3	15
3½	5	2	10½	1½	5½	1½	8½
6	4½	3	13½	8½	6½	3½	18½
5	4½	4	13½	2½	5	3	10½
5½	3	3	11½	8	4	7	19
5	5	1½	11½	5	2	4½	11½
6½	7	4	17½	7	6	6	19
6½	2	1½	10	1	5	1	7
4½	5	2½	12	7½	7	2	16½
4½	3	1½	9	4½	6½	4½	15½
3½	2	2	7½	7½	7½	7	22
5½	4½	3½	13½	1	4	1	6
5½	7	8	20½	2	7	7½	16½
4	4½	½	9	5	3	5	13
4	5	7	16	4½	6	1	11½
5	6	6	17	2	6½	3	11½
1½	4	1½	7	2	2½	5½	10

TABLE 8

Matched Tests Scores for Craigmount High School in Class Groups

Diagram Score	Non Diagram Score	d +	d −	d^2
26	25		1	1
23	22		1	1
26½	24		2½	6.25
21	25	4		16
23	20½		2½	6.25
16½	22½	6		36
16	15½		½	0.25
17½	21	3½		12.25
11½	16½	5		25
24½	26½	2		4
20½	25½	5		25
22	15½		6½	42.25
14½	21	6½		42.25
19	11		8	64
20	18½		1½	2.25
9	14½	5½		30.25
14½	12		2½	6.25
6	12	6		36
6	14	8		64
25½	20		5½	30.25
21	18		3	9

TABLE 8 continued

D Score	ND Score	d +	d −	d^2
16½	16		½	0.25
17½	16		1½	2.25
13½	13½			0
17½	16		1½	2.25
10	15½	5½		30.25
22½	13½		9	81
23	22½		½	0.25
22	20		2	4
23	20		3	9
21	23	2		4
21	20		1	1
20½	14½		6	36
13½	21	7½		56.25
19½	13		6½	42.25
		66½	66	
Σ 644½	645	+ ½		728.25

TABLE 9

Matched Test Scores for ST. Thomas of Aquin's

(Single Class Group)

	D Score	ND Score	d +	d −	d^2
	27	27½	½		0.25
	24½	25½	1		1
	23	22½		½	0.25
	19	20	1		1
	18	18½	½		0.25
	16½	20	3½		12.25
n = 7	16	19½	3½		12.25
			10	½	
Σ	144	153½	+ 9½		27.25

TABLE 10

Matched Test Scores for Kelso High School

(Single Class Group)

	D Score	ND Score	d +	d −	d^2
	25½	27½	2		4
	23	16		7	49
	21½	18½		3	9
	17½	18½	1		1
	21	19½		1½	2.25
	22	20½		1½	2.25
	19½	15		4½	20.25
	17	18	1		1
	17½	11		6½	42.25
	18	16½		1½	2.25
	22	16		6	36
	6½	13½	7		49
	11	12½	1½		2.25
	10½	11½	1		1
	8½	16½	8		64
	9	14½	5½		30.25
	13	9½		3½	12.25
n=19	9	12½	3½		12.25
	5½	16½	11		121
			41½	35	
Σ	297½	304	+ 6½		461.25

TABLE 11

Matched Test Scores for Selkirk High School

	D Score	ND Score	d +	d −	d^2
	24½	26	1½		2.25
	20	18		2	4
	13½	23	9½		90.25
	10	6		4	16
	14½	13½		1	1
	11	9½		1½	2.25
	18½	25	6½		42.25
	12	11		1	1
n = 10	19	16½		2½	6.25
	22	17½		4½	20.25
			17½	16½	
Σ	165	166	+ 1		185.50

TABLE 12

Matched Test Scores for George Heriot's School

	D Score	ND Score	d +	d −	d^2
	11½	11½			0
	10	15	5		25
	16	15		1	1
	17	21½	4½		20.25
	17	22	5		25
	17½	6		11½	132.25
	7½	18½	11		121
	10½	11½	1		1
	12	13	1		1
	14½	19	4½		20.25
	13	19	6		36
	10	8½		1½	2.25
	7	12½	5½		30.25
	9	11½	2½		6.25
	20½	16½		4	16
n=21	11½	16½	5		25
	9	15½	6½		42.25
	13½	15	1½		2.25
	8	10	2		4
	13½	10½		3	9
	13½	7		6½	42.25
Σ	262	295½	61	27½	562.25
			+ 33½		

TABLE 13

Summary of data for the entire sample and calculation of t value.

$\bar{x}_D$	16.446
$\bar{x}_{ND}$	17.000
Σd	51
N	92
Xd	0.554
Σd^2	1964.5
Σx_d^2	1935.7

$$t = \frac{\bar{X}d}{\sqrt{\frac{\Sigma\, x_d^2}{N\,(N-1)}}}$$

$$= \frac{0.554}{\sqrt{\frac{1935.7}{92 \times 91}}}$$

=> t = 1.152 (df 90)

Check Calculations to determine correlation of pretest and post test scores from a sample of Craigmount Pupils.

Pretest Score	Post Test Score			
x	y	xy	x^2	y^2
77	82	6314	5928	6724
74	68	5032	5476	4624
69	68	4692	4761	4624
61	73	4453	3721	5329
53	48	2544	2809	2304
52	63	3276	2704	3969
42	67	2814	1764	4489
38	28	1064	1444	784
Σ 466	497	30189	28607	32847

$$r = \frac{\Sigma xy - \frac{\Sigma x \, \Sigma y}{N}}{\sqrt{\left\{ \Sigma x^2 - \left(\frac{\Sigma x}{N}\right)^2 \right\} \left\{ \Sigma y^2 - \left(\frac{\Sigma y}{N}\right)^2 \right\}}}$$

$$= \frac{158.4}{182.81 \times 246.36}$$

$$\Rightarrow r = 0.73$$

Calculation of Correlation Coefficient

(Diagram effect/Ability Interaction)

Using (D+N) as sum and (D-N) as difference of diagram and non-diagram test scores:-

$$r = \frac{S_{(D+N)\ (D-N)}}{\sqrt{S^2_{(D+N)}\ S^2_{(D-N)}}}$$

where $S_{(D+N)\ (D-N)} = \Sigma(D+N)\ (D-N) - \frac{\Sigma(D+N)\ \Sigma(D-N)}{N}$

$$\overline{\qquad\qquad N \qquad\qquad}$$

and $S^2_{(D+N)} \qquad \Sigma(D+N)^2 - \frac{(\Sigma\ (D+N)\ )^2}{N}$

$$\overline{\qquad\qquad N \qquad\qquad}$$

etc.

$\Sigma\ (D-N) = -505$

$\Sigma\ (D+N) = 30765$

$\Sigma\ (D-N)^2 = 196825$

$\Sigma\ (D+N)^2 = 11127025$

$\Sigma\ (D+N)\ (D-N) = -\ 107825$

$N = 92$

$$r = \frac{61048.1}{194052.99 \times 839142.12}$$

$$\Rightarrow \quad r = \underline{0.151}$$

Fisher's Transformation for estimation of significance of a correlation coefficient.

$$Z = \tfrac{1}{2} \log_e \frac{1 + r}{1 - r}$$

$$= 1.15 \log_{10} \frac{1 + r}{1 - r}$$

$$= \underline{0.152}$$

$$\text{deviation } z = \frac{Z}{S_Z} = \frac{0.152}{0.106}$$

$$\Rightarrow \quad \underline{z = 1.44}$$

Calculation of Correlation Coefficient for matching of D and ND Pairs.

$$r = \frac{S_{D,N}}{\sqrt{S_{DD} \quad S_{NN}}}$$

Σ D	=	1513.5
Σ N	=	1563.5
Σ D^2	=	2777050
Σ N^2	=	2884875
Σ DN	=	2732550

=> r = 0.965

Means of 3 Ability Groups in Tests

	D	ND
TOP	20.71	20.47
MIDDLE	16.31	16.47
LOWER	12.55	14.31

GRAPH OF D, AND ND MEANS

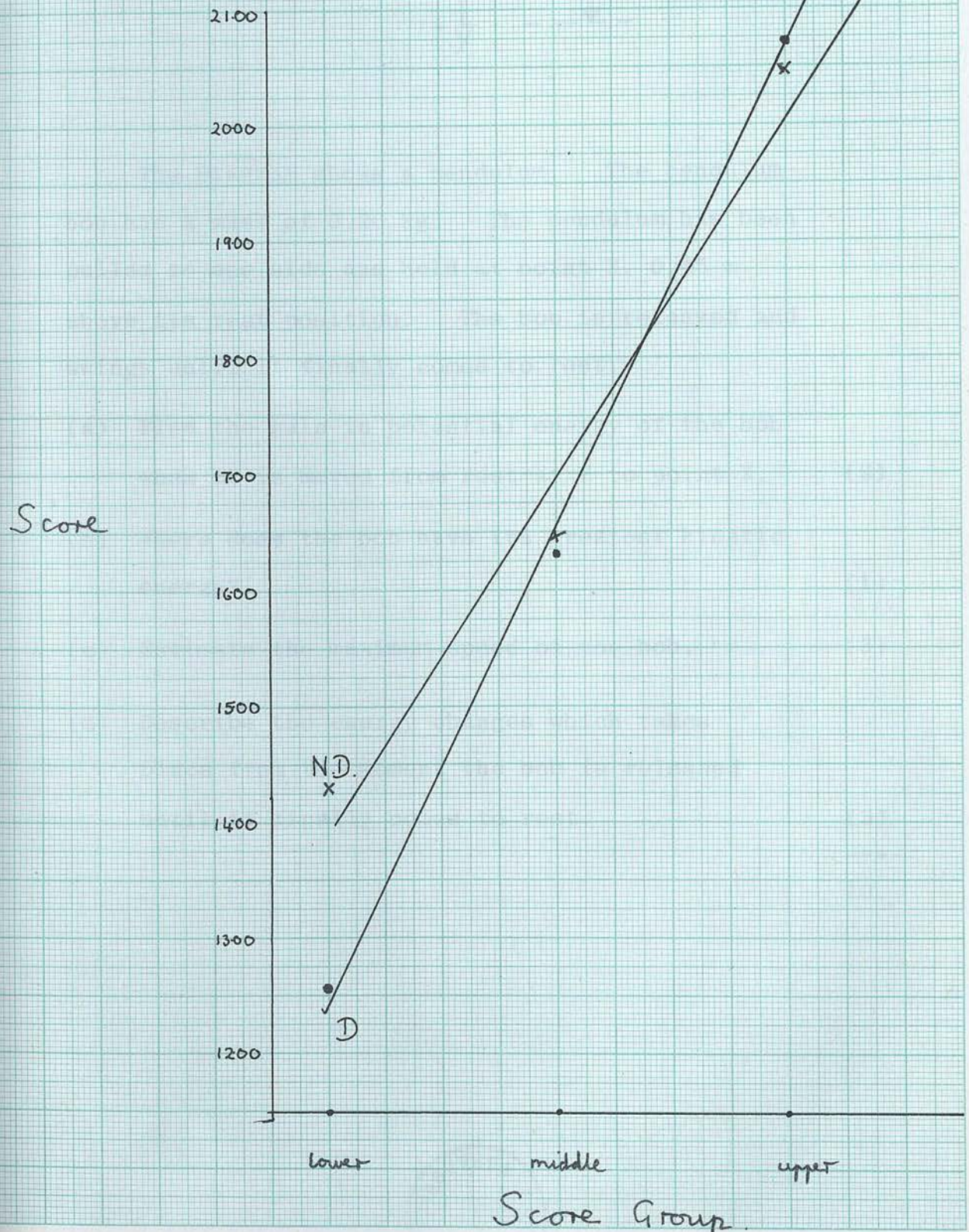
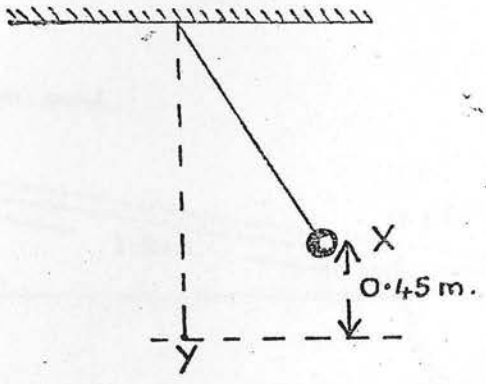


DIAGRAM TEST QUESTION



The Diagram shows a pendulum. The pendulum bob has a mass of 0.20 kg. The pendulum has been pulled to one side and held at point X. 0.45 m. above its rest position. The bob is released and swings until it finally comes to rest.

(a) Find the gain in potential energy of the bob when it is moved from rest at Y to point X. (3)

(b) Where does the bob have its greatest kinetic energy? (1)

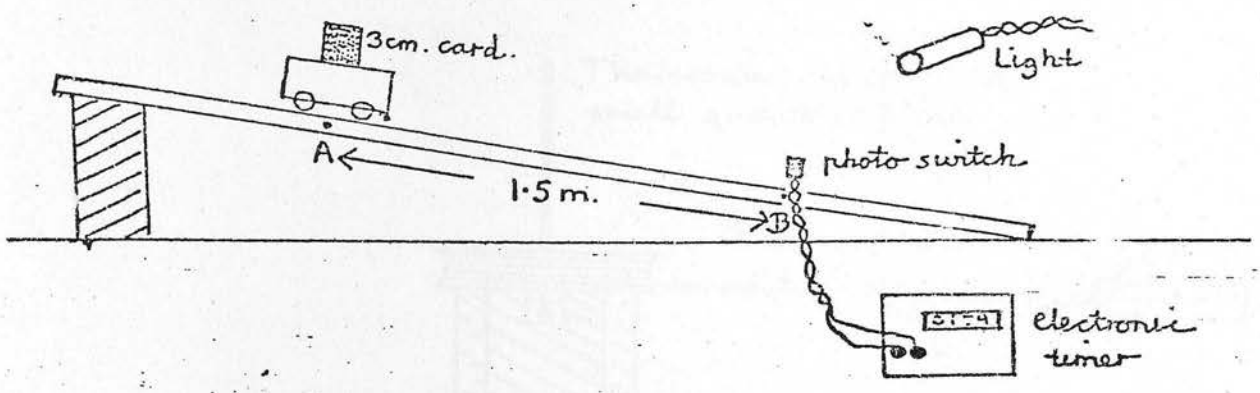
(c) Estimate the maximum speed of the bob. (3)

(d) Describe the energy changes which takes place from the moment the bob is released until it finally comes to rest. (3)

10.

DIAGRAM TEST QUESTION

2.



The trolley was released from rest at A and reached B 5.0s later as timed on a stop-watch. As the trolley passed B, the card interrupted the light beam and the timer registered 0.05s. The timer was accurate to one millisecond.

(a) Find:- (i) the acceleration of the trolley (3)

(ii) the average velocity between A and B (2)

(b) Give two reasons why stopwatches could not be used for both time measurements. (2)

(c) What is the effect on the acceleration of:-

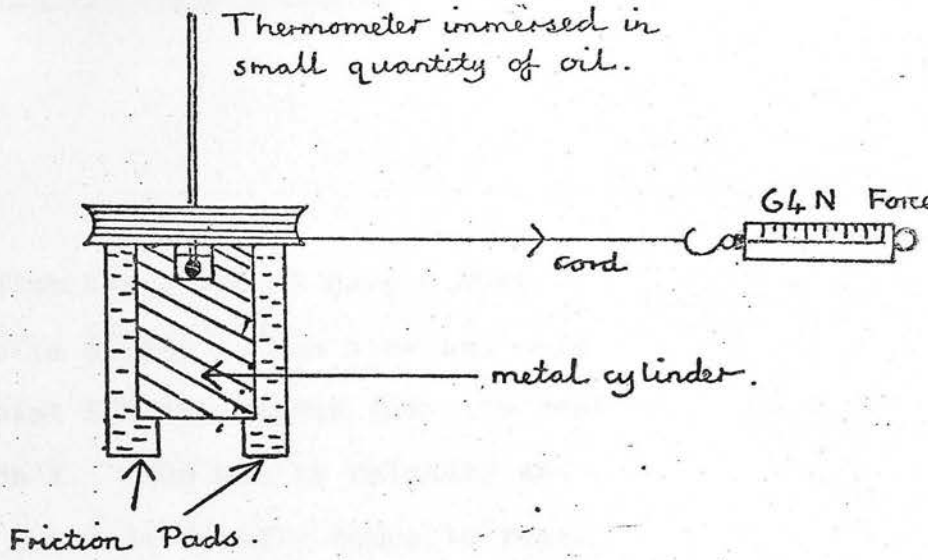
(i) increasing the tilt of the track.

(ii) increasing the mass of the trolley.

(iii) starting further up the track? (3)

10

DIAGRAM TEST QUESTION



A metal cylinder of mass 0.3kg. is rotated between two friction pads by pulling a cord with a steady force of 64N. When the cord is pulled 7.5m a temperature rise of 4K is produced in the cylinder.

(a) Assuming no heat losses, calculate the specific heat capacity of the metal. (5)

(b) Why is the thermometer immersed in oil? (1)

(c) (i) Suggest two ways in which a higher rise in temperature could be produced in the metal cylinder in this experiment. (2)

(ii) Discuss whether a higher rise in temperature would lead to a more accurate value for the specific heat capacity of the metal. (2)

10

NON-DIAGRAM TEST QUESTION

1. A pendulum has a bob of mass 0.20kg. The bob is pulled to one side and held at a point X 0.45m higher than its rest position Y. The bob is released and swings until it finally comes to rest.

   (a) Find the gain in potential energy of the bob when it is moved from rest at Y to point X. (3)

   (b) Where does the bob have its greatest kinetic energy? (1)

   (c) Estimate the maximum speed of the bob. (3)

   (d) Describe the energy changes which take place from the moment the bob is released until it finally comes to rest. (3)

   10

NON-DIAGRAM TEST QUESTION

2. A piece of card 3cm long was attached to a trolley on a tilted slope. The trolley was released from rest and travelled 1.5m down the slope in 5.0s as timed on a stopwatch. At this point, the card interrupted a light beam and photo switch arrangement connected to an electronic timer. The timer which was accurate to one millesecond registered 0.05 s.

(a) Find:- (i) the acceleration of the trolley (3)

(ii) the average velocity over 1.5m (2)

(b) Give two reasons why stopwatches could not be used for both time measurements. (2)

(c) What is the effect on the acceleration of:-

(i) increasing the tilt of the track

(ii) increasing the mass of the trolley

(iii) starting further up the track ? (3)

10

NON-DIAGRAM TEST QUESTION

3. A metal cylinder of mass 0.3kg. is rotated between two friction pads by pulling a cord with a steady force of 64N. When the cord is pulled 7.5 m, a temperature rise of 4K is produced in the cylinder. The rise in temperature is recorded on a thermometer placed in a hole in the cylinder and surrounded by a small quantity of oil.

(a) Assuming no heat losses, calculate the specific heat capacity of the metal. (5)

(b) Why is the thermometer immersed in oil? (1)

(c) (i) Suggest two ways in which a higher rise in temperature could be produced in the metal cylinder in this experiment. (2)

(ii) Discuss whether a higher rise in temperature would lead to a more accurate value for the specific heat capacity of the metal. (2)

10

REFERENCES

1. Varley, W.H., Levin, J.R., Severson, R.A. and Wolff, P.
Training imagery production in young children through motor involvement.
Journal of Educational Psychology 66 : 262-266 1974.

2. Moser, C.A.
Survey Methods in Social Investigation
Heinemann 1965 pps 121-127.

3. Pilliner, A.E.G. in
Experiment in Educational Research
Educational Studies: A third level course Block 5.
Open Univiversity Press.

4. Campbell, D.T. and Stanley, J.C.
Experimental and Quasi-Experimental Designs for Research in Teaching.
In Gage, N.L. (Ed.,) Handbook for Research on Teaching
Rand McNally and Co. 1964 pps 175-176.

5. Guilford, J.P. and Fruchter, B.
Fundamental Statistics in Psychology and Education.
McGraw-Hill Kogakusha 1978. pps 157-158

6. Johnson, M.C.
A Review of Research Methods in Education.
Rand McNally. 1977. Ch. 7.

7. Scottish Certificate of Education Examination Board.
Physics. Ordinary and Higher Grades - Syllabuses and Notes.
S.C.E. Examination Board 1969 pps 28-43.

8. Scottish Education Department
Curriculum Paper No. 7 - Science for General Education.
H.M.S.O. 1969 p27.

GENERAL REFERENCES

MORRIS, L. L., FITZ-GIBBON, C.T.

HOW TO PRESENT AN EVALUATION REPORT

Sage Publications 1978

EVANS, K.M.

PLANNING SMALL-SCALE RESEARCH

N.F.E.R 1978

A1	Education in Science	1978	-	1980
A2	Research in Education	1968	-	1980
A3	Educational Research	1954	-	1980
A4	British Journal of Education Psychology	1954	-	1980
A5	Physics Education	1954	-	1980